Taking the data together—that is, the original seismic data[13], the borehole data[14], the large concentration increases in the magmatic gases CO_2 and He, and the increase in ^{3}He/heat ratio—we conclude that the 8 June 1999 seismic swarm on Endeavour was the result of magma movement beneath the ridge crest, and not of a purely tectonic event. The data presented here confirm that magmatic events have profound effects on the characteristics of hydrothermal fluids. Magmatic activity causes transient fluxes of magmatic gases and shifts of pressure, temperature and redox conditions in the high-temperature reaction zone, thereby influencing fluid phase separation and causing large changes in the Cl concentration of hydrothermal fluids. Magmatic events can cause significant changes in the flux of ore-forming metals to the sea floor[27,30]. The concentrations of CO_2, He and H_2 rise rapidly during magmatic events, but these concentration changes are not yet well constrained. The volatile flux during the first few months after magmatic events clearly needs to be better evaluated, as the resulting contribution of volatiles during this early period may rival that released by a mature hydrothermal system during an entire year. □

Received 2 December 2002; accepted 5 March 2003; doi:10.1038/nature01569.

1. Kelley, D. S., Baross, J. A. & Delaney, J. R. Volcanoes, fluids, and life at mid-ocean ridge spreading centers. *Annu. Rev. Earth Planet. Sci.* **30,** 385–491 (2002).
2. Delaney, J. R., Robigou, V., McDuff, R. E. & Tivey, M. K. Geology of a vigorous hydrothermal system on the Endeavour segment, Juan de Fuca Ridge. *J. Geophys. Res.* **971,** 19663–19682 (1992).
3. Butterfield, D. A. *et al.* Gradients in the composition of hydrothermal fluids from the Endeavour segment vent field: phase separation and brine loss. *J. Geophys. Res.* **99,** 9561–9583 (1994).
4. Johnson, H. P. *et al.* Earthquake-induced changes in a hydrothermal system on the Juan de Fuca mid-ocean ridge. *Nature* **407,** 174–177 (2000).
5. Kappel, E. S. & Ryan, W. B. F. Volcanic episodicity and a non-steady state rift valley along northeast Pacific spreading centers: evidence from Sea MARC I. *J. Geophys. Res.* **91,** 13925–13940 (1986).
6. Rohr, K. M. M., Milkereit, B. & Yorath, C. J. Asymmetric deep crustal structure across the Juan de Fuca Ridge. *Geology* **16,** 533–537 (1988).
7. White, D. J. & Clowes, R. M. Shallow crustal structure beneath the Juan de Fuca Ridge from 2-D seismic refraction tomography. *Geophys. J. Int.* **100,** 349–367 (1990).
8. Wilcock, W. S. D., Archer, S. D. & Purdy, G. M. Microearthquakes on the Endeavour segment of the Juan de Fuca Ridge. *J. Geophys. Res. B* **107,** doi:101029/2001JB000505 (2002).
9. Lister, C. R. B. in *Hydrothermal Processes at Seafloor Spreading Centers* (eds Rona, P. A., Boström, K., Laubier, L. & Smith, K. L. Jr) 141–168 (Plenum, New York, 1983).
10. Wilcock, W. S. D. & Delaney, J. R. Mid-ocean ridge sulfide deposits: evidence for heat extraction from magma chambers or cracking fronts? *Earth Planet. Sci. Lett.* **145,** 49–64 (1996).
11. Detrick, R. S. *et al.* New multichannel seismic constraints on the crustal structure of the Endeavour Segment, Juan de Fuca Ridge: Evidence for a crustal magma chamber. *Eos* **83,** F1353 (2002).
12. Lowell, R. P. & Germanovich, L. N. On the temporal evolution of high-temperature hydrothermal systems at ocean ridge crests. *J. Geophys. Res.* **99,** 565–575 (1994).
13. Bohnenstiehl, D. R., Tolstoy, M., Dziak, R. P., Fox, C. G. & Smith, D. K. Aftershock sequences in the mid-ocean ridge environment: an analysis using hydroacoustic data. *Tectonophysics* **354,** 49–70 (2002).
14. Davis, E. E., Wang, K., Thompson, R. E., Becker, K. & Cassidy, J. F. An episode of seafloor spreading and associated plate deformation inferred from crustal fluid pressure transients. *J. Geophys. Res.* **106,** 21953–21963 (2001).
15. Butterfield, D. A., Massoth, G. J., McDuff, R. E., Lupton, J. E. & Lilley, M. D. Geochemistry of hydrothermal fluids from Axial Seamount hydrothermal emissions study vent field, Juan de Fuca Ridge: subseafloor boiling and subsequent fluid-rock interaction. *J. Geophys. Res.* **95,** 12895–12921 (1990).
16. Sedwick, P. N., McMurtry, G. M. & MacDougall, J. D. Chemistry of hydrothermal solutions from Pele's Vents, Loihi seamount, Hawaii. *Geochim. Cosmochim. Acta* **56,** 3643–3667 (1992).
17. Lupton, J. E., Lilley, M. D., Olson, E. J. & Von Damm, K. L. Gas chemistry of vent fluids from 9°–10° N on the East Pacific Rise. *Eos* **72,** F481 (1991).
18. Lilley, M. D., Olson, E. J., McLaughlin, E. & Von Damm, K. L. Methane, hydrogen and carbon dioxide in vent fluids from the 9°N hydrothermal system. *Eos* **72,** F481 (1991).
19. Dixon, J. E., Stolper, E. M. & Holloway, J. R. An experimental study of water and carbon dioxide solubilities in mid-ocean ridge basaltic liquids. Part I: Calibration and solubility models. *J. Petrol.* **36,** 1607–1631 (1995).
20. Dixon, J. E., Stolper, E. & Delaney, J. R. Infrared spectroscopic measurements of CO_2 and H_2O in Juan de Fuca Ridge basaltic glasses. *Earth Planet. Sci. Lett.* **90,** 87–104 (1988).
21. Kadko, D. & Butterfield, D. A. The relationship of hydrothermal fluid composition and crustal residence time to maturity of vent fields on the Juan de Fuca Ridge. *Geochim. Cosmochim. Acta* **62,** 1521–1533 (1998).
22. Stuart, F. M. & Turner, G. Mantle-derived ^{40}Ar in mid-ocean ridge hydrothermal fluids: implications for the source of volatile and mantle degassing rates. *Chem. Geol.* **147,** 77–88 (1998).
23. Seewald, J. S., Cruse, A. M. & Saccocia, P. J. Aqueous volatiles in hydrothermal fluids from the Main Endeavour vent field: Temporal variability following earthquake activity. *Eos* **82,** F615 (2001).
24. Von Damm, K. L. *et al.* Evolution of East Pacific Rise hydrothermal vent fluids following a volcanic eruption. *Nature* **375,** 47–50 (1995).
25. Sohn, R. A., Fornari, D. J., Von Damm, K. L., Hildebrand, J. A. & Webb, S. C. Seismic and hydrothermal evidence for a cracking event on the East Pacific Rise crest at 9° 50′ N. *Nature* **396,** 159–161 (1998).
26. Lupton, J. E., Baker, E. T. & Massoth, G. J. Helium, heat, and the generation of hydrothermal event plumes at mid-ocean ridges. *Earth Planet. Sci. Lett.* **171,** 343–350 (1999).
27. Seyfried, W. E. Jr & Ding, K. *Seafloor Hydrothermal Systems: Physical, Chemical, Biological and Geological Interactions* (eds Humphris, S. E., Zierenberg, R. A., Mullineaux, L. S. & Thomson, R. E.) 248–272 (American Geophysical Union, Washington DC, 1995).
28. Ding, K., Seyfried, W. E. Jr, Tivey, M. K. & Bradley, A. M. In situ measurement of dissolved H_2 and H_2S in high-temperature hydrothermal vent fluids at the Main Endeavour Field, Juan de Fuca Ridge. *Earth Planet. Sci. Lett.* **186,** 417–425 (2001).
29. Lilley, M. D. *et al.* Anomalous CH_4 and NH_4^+ concentrations at an unsedimented mid-ocean-ridge hydrothermal system. *Nature* **364,** 45–47 (1993).
30. Von Damm, K. L. Chemistry of hydrothermal vent fluids from 9° - 10° N, East Pacific Rise: "Time zero," the immediate posteruptive period. *J. Geophys. Res.* **105,** 11203–11222 (2000).

Supplementary Information accompanies the paper on *Nature*'s website (**http://www.nature.com/nature**).

Acknowledgements We thank L. Evans and K. Roe for technical assistance, D. Kelley for a critical reading of the manuscript, W. Wilcock for discussions, and W. E. Seyfried Jr for the opportunity to join his 1999 cruise to Endeavour. The manuscript was improved by comments and suggestions from J. Seewald and R. Lowell. This work was supported by the National Science Foundation and in part by the National Oceanic and Atmospheric Administration VENTS Program.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.D.L. (e-mail: lilley@ocean.washington.edu).

A test of the unified neutral theory of biodiversity

Brian J. McGill

Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721, USA

One of the fundamental questions of ecology is what controls biodiversity. Recent theory suggests that biodiversity is controlled predominantly by neutral drift of species abundances[1–4]. This theory has generated considerable controversy[5–12], because it claims that many mechanisms that have long been studied by ecologists (such as niches) have little involvement in structuring communities. The theory predicts that the species abundance distribution within a community should follow a zero-sum multinomial distribution (ZSM), but this has not, so far, been rigorously tested. Specifically, it remains to be shown that the ZSM fits the data significantly better than reasonable null models. Here I test whether the ZSM fits several empirical data sets better than the lognormal distribution. It does not. Not only does the ZSM fail to fit empirical data better than the lognormal distribution 95% of the time, it also fails to fit empirical data better even a majority of the time. This means that there is no evidence that the ZSM predicts abundances better than the much more parsimonious null hypothesis.

The unified neutral theory of biodiversity (or UNTB, hereafter used to refer to either the theory or Hubbell's 2001 book[1]) predicts that species abundance distributions (SADs) should follow the ZSM (see Box 1). To test whether the ZSM fits the data better than the null lognormal hypothesis, I used two data sets. First, I used the North American Breeding Bird Survey[13–15] (BBS), for the simple reason that it is one of the few data sets that has replicates (in this case different sites) taken with identical sampling methodologies. Replicates are important for testing statistical significance. Since the ZSM has been applied primarily to intensively sampled data, I averaged the BBS data over a five-year period (1996–2000). Thus, rare species that show up only once in five years are included. I took 100 replicates by randomly selecting 100 routes that were rated to be of high quality for all five years. As an alternative data source, I also used the well-known Barro Colorado Island (BCI) tree data set, in

which trees are sampled in a 50-ha plot in Panama[10,16]. This has the disadvantage of providing only a single data set, and hence not allowing an analysis of replicates. However, it is much larger (number of individuals and species) than a given BBS route, and is exhaustively sampled. Most importantly, it has been used extensively in discussions on SADs and the ZSM[1,2,10].

The lognormal and ZSM distributions were fitted to the data using maximum-likelihood methods. Fitting the lognormal distribution to the data is quite simple; one merely log-transforms the data and then uses the usual methods for fitting the normal distribution (using the sample mean and a bias-corrected sample variance). Details of the estimating methods used for the ZSM can be found in Box 1 and the Supplementary Information. Source code is available from the author at ⟨http://www.brianmcgill.org/zsmcode.html⟩. Eight measures of goodness-of-fit were then calculated and compared (described in the legends of Tables 1–3). Fits for the truncated lognormal were also calculated, but were found to perform slightly worse than the lognormal (although better than the ZSM in general), and are not presented. If the ZSM is better than the lognormal in 50% of cases (50 routes), then it performs better than the null. If it is better 95% of the time (95 out of 100 routes), then it is statistically significantly better.

Several issues were encountered while developing the fitting routines for the ZSM that have scientific interest and are described here: (1) Time to equilibrium. The ZSM fitting routine generates a 'metacommunity' population of very large size, takes a sample of the size of the local community, and then iterates the appropriate equations (see page 86 of ref. 1) for a number of time steps (individual deaths). No indication of what fixed number of time steps is appropriate has been published. My own experiments show that it can be large. For a local community of 1,600 individuals, it can take 10^7 time steps (individual deaths) to reach equilibrium. For a local community of 10,000 individuals, it can take 10^6 individual deaths (Fig. 1). Assuming a tree mortality rate of 1% per year (most studies find 1–2% per year under normal conditions[17,18]), 10^6 time steps translates into 10,000 years just to reach equilibrium. Even

Box 1
Theory, testing and estimation of the ZSM

The SAD describes the relative abundances of different species within a community (see Fig. 3 for two ways to plot this information). The SAD nearly always shows many rare species and a few highly abundant species. Well over two dozen different probability distributions have been suggested as the 'right' SAD, with more appearing every year[1,20–23]. The UNTB[1] has proposed that the SAD should follow a ZSM distribution.

The UNTB predicts that the species diversity and relative abundances of species can be explained by the neutral drift of the abundances of different species, in direct analogy to the neutral theory of molecular evolution[24]. In particular, the UNTB suggests that the number of individuals in a community, J_M, is constant, and that at each time step, one random individual dies and is replaced either by a new species with probability ν, or by an offspring of one of the randomly selected remaining individuals with probability $1 - \nu$. Then the distribution of abundances (SAD) should be determined by the compound parameter $\theta = 2J_M\nu$, known as the fundamental biodiversity number. Hubbell calls this the 'metacommunity', and adds one further hierarchical level, the local community. In the local community, of population size J (much smaller than J_M, but to an unspecified extent), speciation does not occur, but when an individual dies it is replaced by a random sample from the metacommunity with probability m, or within the local community with probability $1 - m$, where m is the migration probability. Technically, this lower hierarchical level is a modified ZSM, but UNTB calls it just the ZSM, and I will follow this convention. Thus, the UNTB predicts that the SAD will follow the ZSM, which is parameterized by three numbers: θ, J and m.

The strongest test of this hypothesis to date is in Hubbell's recent book (referred to as UNTB here)[1], which has provided a number of examples showing that the ZSM fits empirical data. However, the method by which goodness-of-fit of the ZSM to the data is measured is by visual examination (see, for example, Fig. 5.9 of the UNTB[1]). No objective measure of goodness-of-fit is reported. Usually, science demands a stronger test, such as the data fitting well by some objective measure (for example, r^2) or preferably fitting better or even statistically significantly better than some null hypothesis[25]. A reasonable null hypothesis for SADs is the lognormal distribution. Owing to arguments based on the central limit theorem[26], there are good statistical, non-biological reasons to expect the lognormal distribution to fit SADs[27]. In this paper, I test whether the ZSM fits empirical data statistically significantly better than the lognormal.

The ZSM does not have an analytical form, so we can't use likelihood ratios or other analytical methods of significance[28]. For this reason I use a data set with replicates (the BBS) to allow significance testing. Moreover, fitting the ZSM distribution is quite complicated. I implement the method described in UNTB[1] (pages 289–294). Details over and above Hubbell's original description are found in the Supplementary Information. Briefly, the method uses a generator function to generate one random sample from the ZSM. If we generate many (for example, 100) samples, and average across these samples, then we have a good estimate of the ZSM for a given set of parameters. We can evaluate the fit for different parameters using a generic likelihood function, and then search for the two parameters θ and m, which maximizes the observed likelihood. The algorithm was written in C and Matlab, and highly tuned for performance. Despite this, it takes half an hour of computation on an Athlon 500-MHz machine to fit the BCI data. A plot of the calculated likelihood surface can be found in the Supplementary Information. In view of the computational intensity of this calculation, we use a very simple hierarchical iterative method to search for the maximum likelihood, which is adequate for this problem because of the relatively smooth nature of the likelihood surface.

Table 1 **Goodness-of-fit measures for BBS data**

	r^2	r^2 MC	r^2 corr.	K
Lognormal	1.00 (0.99,1)	0.98 (0.93,0.99)	0.98 (0.96,1)	0.10 (0.06,0.15)
ZSM	0.99 (0.95,1)	0.89 (0.67,0.97)	0.97 (0.92,0.99)	0.26 (0.20,0.37)
ZSM beats lognormal	4%	4%	23%	0%

Compares goodness-of-fit for lognormal and ZSM distributions for 100 routes from the BBS. In each cell, the first number represents the mean across 100 routes; the numbers in parentheses represent the 2.5 and 97.5 percentiles (similar to a 95% confidence interval). The last row indicates the percentage of routes out of 100 where ZSM scores better than the lognormal. These four measures of goodness-of-fit are all based on calculating the observed cumulative distribution function (CDF), as is done in the Kolmogorov–Smirnov test. This is then compared with the theoretically predicted CDF using three measures of r^2—raw r^2, mean-corrected r^2 (r^2 MC) and square of the Pearson correlation r (r^2 corr.)—and the traditional Kolmogorov–Smirnov statistic, K, measuring greatest absolute vertical deviance.

Table 2 **χ^2 goodness-of-fit measures for BBS data**

	χ^2	χ^2 Preston $\log_2$ bins	χ^2 10 + 1 bins	χ^2 5 bins
Lognormal	19.7 (5.2,63.7)	8.97 (2.71,17.60)	11.1 (3.6,22.9)	8.64 (1.27,34.74)
ZSM	Inf. (15.9,NaN)	Inf. (12.912,Inf.)	19.0 (7.6,NaN)	Inf. (5.949,NaN)
ZSM ≠ INF	7	96	95	47
ZSM	20.7	24	19	12.9
ZSM beats lognormal	28.6%	0%	3.16%	6.38%

As in Table 1, but for various χ^2 measures of goodness-of-fit based on various bin schemes. The χ^2 statistic is undefined when the expected number of occurrences in a bin is zero. This causes the Inf. (infinite or x/0) and NaN (0/0) results. Although it biases results towards the ZSM, I also report the results for cases where Inf. and NaN results are thrown out. The number of remaining cases out of 100 is reported in the third row, and the average χ^2 value for just these cases is reported in the fourth row; the percentage of times the ZSM beats the lognormal is summarized in the last row. The four binning schemes, from left to right, are: (1) divide the observed range into 10 bins on an arithmetic scale; (2) divide the observed range into bins on a logarithmic scale according to Preston's rules; (3) divide the lower 80% of the data into 10 arithmetic bins and include the upper 20% in one bin; and (4) similar to method 1 but with only 5 bins. Methods (2)–(4) are all attempts to ensure that the right-most bins are populated despite data that include very few species with very high abundances.

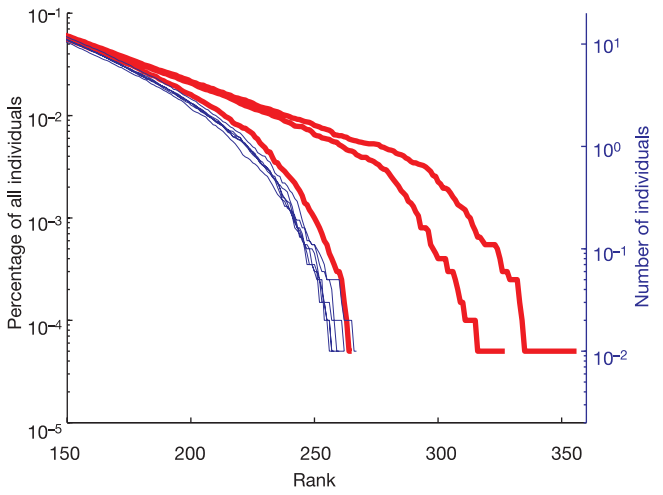

Figure 1 Approach to equilibrium of a local community. This is the lower right corner of a standard rank-abundance plot (see Fig. 3a for a full rank-abundance plot) with abundance on the vertical axis and rank on the horizontal axis (each point is one species; ranks 1–150 were omitted because they were identical for all nine lines; rank 1 is assigned to the most abundant species). The three red lines represent the state of the local community as it approaches equilibrium for 1×10^3, 1×10^4 and 1×10^5 time steps (that is, individual deaths and going from right to left, respectively). We can see that by 1×10^5 deaths, the system is close to equilibrium, but is still noticeably different (at least over the range of rank 200–270). The right-most red line is closest to the source metacommunity (although the actual theoretical metacommunity would continue as a straight line rather than curving down). The six blue lines represent 1×10^6, 2×10^6, 3×10^6, 4×10^6, 5×10^6 and 1×10^7 time steps that have occurred in the local community since it became partially isolated from the metacommunity. The blue lines are all interpreted as being at equilibrium, as the 1×10^6 line is left-most, the 2×10^6 line is right-most, and the 1×10^7 line is in the centre. Within this noisy equilibrium, diversity drops down to about 250 species, jumps up to about 270 species for the next recorded time period (2×10^6 deaths), and oscillates in between for the rest of the time. This simulation had $\theta = 50$, $m = 0.1$ and $J = 20{,}000$—close to the BCI data set.

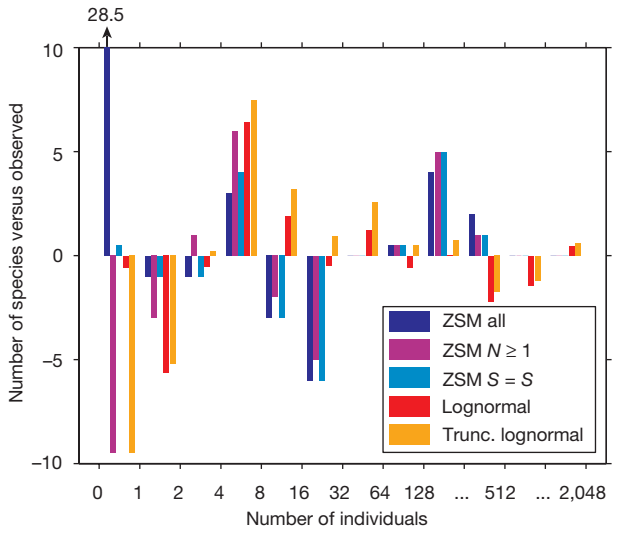

Figure 2 Deviations from observed abundances depends on the stopping rule. This figure shows deviations from observed abundances by Preston $\log_2$ bins for various stopping rules on the ZSM and lognormal. A positive value indicates predicting more species than observed. The vertical axis is truncated as the left-most bar should actually go up to 28.5. Trunc., truncated.

assuming that a local community has 100,000 individuals, it would still take about 1,000 years. Thus, the time that the theory predicts it will take to reach local equilibrium may in fact be longer than the time over which environmental conditions are constant enough to support a given equilibrium. (2) Noisy equilibrium. The local community equilibrium is a very noisy one. Species diversities fluctuate over a range of about 10% of total species diversity (Fig. 1). This does not mean that the UNTB is not an equilibrium theory, but it is important to realize just how variable the equilibrium is. (3) Number of parameters and stopping rules. The methods used for generating the ZSM produce a large number of very rare species. In all of our runs, species were produced with abundances down to 0.01 individuals. This produced poor fits to the empirical data (too many rare species). The obvious solution is to truncate or stop the distribution somewhere. Looking at the figures in the UNTB, it is clear that the method used is to stop the distribution at the same number of species as the observed data. It would also have been reasonable to eliminate species with an abundance less than one. This (or any other stopping method) adds a third, undocumented fitting parameter, which makes the ZSM less parsimonious than it first appears to be.

This issue of the stopping method is pivotal in one of the main claims to success of the ZSM—predicting the abundances of rare species better than the lognormal (see figures 5.7 and 5.8, and the associated discussion, in the UNTB[1]). The comparison shown in the two figures is based on having the ZSM stop at the observed number of species. If the ZSM is stopped where abundance is ≥ 1, then it underpredicts the number of species. If the ZSM is not stopped, then it badly overpredicts the number of rare species. When the lognormal is stopped where abundance is ≥ 1, it is called the truncated lognormal. As mentioned above, my own tests show that the truncated lognormal fits worse than the lognormal (presumably because the truncated lognormal is right-skewed, whereas exhaustively sampled abundance data is left-skewed[19]). In figures 5.7 and 5.8 of the UNTB[1], Hubbell compares the stopping rule for the ZSM that causes it to have the best possible fit with the stopping rule for the lognormal that causes the worst possible fit (that is, the truncated lognormal; see Fig. 2). When I compare the best ZSM stopping rule with the best lognormal stopping rule (Fig. 2), differences are small. Moreover, differences are also small, with the ZSM sometimes performing worse, if we make parallel comparisons of untruncated ZSM with untruncated lognormal, or of truncated ZSM with truncated lognormal. This sensitive dependency of the ZSM fit on the stopping rule means that we must examine carefully the statistical and scientific validity of stopping on the basis of the number of species. In the rest of this paper, I use the stopping rule giving the best fit in each case (having the same number of species as observed data for the ZSM as was done by Hubbell, and being untruncated for the lognormal).

The results of comparing goodness-of-fit for the BBS bird data

Table 3 **Goodness-of-fit measures for the BCI tropical tree data**

	r^2	r^2 MC	r^2 corr.	K	χ^2	Preston log χ^2	χ^2 10 + 1 bins	χ^2 5 bins
Lognormal	1	0.996	0.997	0.0565	12.7	5.45	10.3	4.57
ZSM	0.998	0.979	0.994	0.154	NaN	24.6	22.8	Inf.
ZSM beats lognormal	No	No	No	No	NA	No	No	No

Goodness-of-fit measures are as in the previous two tables. Only one data set is used, so no confidence intervals are presented. The last row indicates whether the ZSM scores better than the lognormal. NA, not applicable.

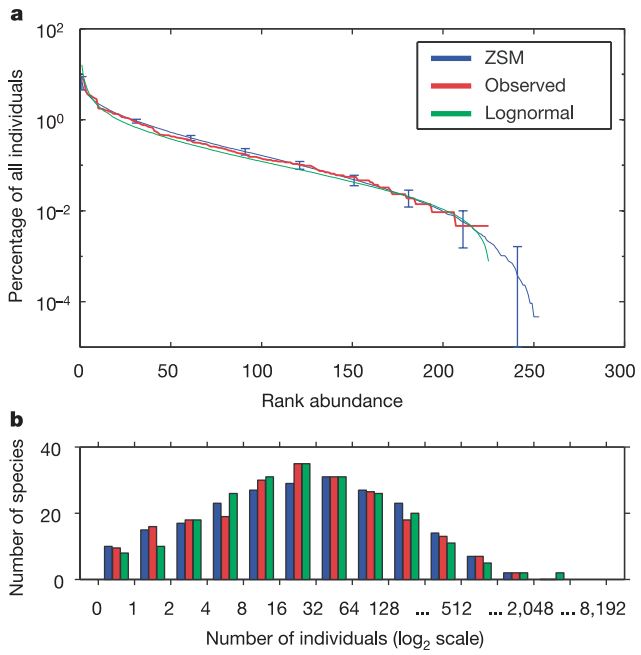

Figure 3 Comparing the fitted and actual BCI data. **a**, A standard rank-abundance plot (with abundance on the vertical axis and rank on the horizontal axis) of the estimated ZSM distribution versus the lognormal distribution versus actual data from the BCI data set. The error bars give one standard deviation above and below the fitted curve for the ZSM. **b**, A binned histogram using the same colour scheme as in **a** and with binning on a $\log_2$ scale. Thus, the bars between the tick marks for 2 and 4 indicate how many species with two to four individuals were calculated. Species with exactly two individuals are split between the 1–2 and 2–4 bins and so forth for all bin boundaries.

are summarized in Tables 1 and 2, and for the BCI tree data in Table 3. The ZSM performs reasonably well if the only criterion is to provide a good fit. However, when compared with the null hypothesis of the lognormal, the ZSM fares poorly. The average fits for the ZSM for all three types of r^2 are lower than the lognormal, and the ZSM beats the lognormal on only about 5–25% of the cases. It does much worse on the K (Kolmogorov–Smirnov) statistic, with an average K of 0.26 compared with 0.10 for the lognormal, and beats the lognormal in no case. Interpreting the χ^2 results is more complicated. The χ^2 statistic is undefined when the expected number of observations is zero. Because the ZSM distribution can be calculated only approximately, it regularly produces expected counts of zero. This results in either infinite χ^2 ($x/0$) or undefined χ^2 (0/0). Although such cases (especially the infinite) should probably count against the ZSM (they do not occur for the lognormal), I threw out all such cases and then compared the results with the lognormal for the remaining cases. Although this should bias the results towards the ZSM, the ZSM performs much worse than the lognormal on all χ^2 measures.

The 100 BBS routes averaged 77.6 species per route (with a 95% range of 32–105). These same routes averaged 3,875 individuals (over five years) with a 95% range of 1,325–7,023. The parameters for the ZSM on this data were as follows: the fundamental biodiversity number, θ, averaged 18.46 with a 95% range of 5.9–31.9, whereas migration probability, m, averaged 0.38 with a 95% range of 0.048–1. Interestingly, there was a strong correlation between species diversity, S, and the fundamental biodiversity number ($\theta = -3.94 + 0.289S$, $P < 0.0001$, $n = 100$, $r^2 = 0.48$; visual analysis of residuals indicates that untransformed linear regression is appropriate).

The results for the BCI tree data set are similar, although we cannot proceed to replicated comparisons against a null hypothesis, because we have only one data set. The lognormal beats the ZSM on all measures of goodness-of-fit. Our estimated parameters were $\theta = 48.5$ and $m = 0.079$, resulting in an estimated curve that gave a good fit to the data (Fig. 3). My estimated parameters are close to the UNTB's estimates[1] of $\theta = 50$ and $m = 0.1$, and the difference is presumably due to fitting the BCI data for a year or a cutoff in minimum tree size that is different from that used by Hubbell (the number of individuals is very different from what Hubbell reports). Nonetheless, to make sure that this did not cause a poor fit, I also calculated goodness-of-fit using Hubbell's estimates of $\theta = 50$ and $m = 0.1$, and it did not change the results materially (usually only in the third significant digit).

A few cautions are in order when interpreting these results. First, it must be concluded that estimating the parameters for the ZSM is still an art and not an exact process, so there may be some room for improvement. However, given that I explored many approaches and always chose the one that made the ZSM perform best, I suspect that any further improvements will be small compared with the degree to which it falls short of the lognormal. Second, it should be noted that the UNTB also makes predictions regarding species area curves and phylogenies, which I have not addressed. Finally, it is important to note that the success of the lognormal does not mean that community structure is random; it simply means that community structure is a function of several multiplicative processes (the number of processes need not be large for the central limit theorem to produce a shape close to lognormal).

This paper has shown that the central empirical test of the UNTB (superior fit of the ZSM to observed abundance data; see chapter 9 of the UNTB[1]) is in fact not true. The ZSM does indeed fit well. However, given the fact that the lognormal performs better, we cannot give the UNTB any special priority over the two dozen other theories that fit abundance data reasonably well. It is also important to note the extent of the failure to fit better than the lognormal. By using a data set with consistent methods (BBS) and many sites, we were able to quantify the relative performance; the lognormal outperforms the ZSM roughly 95–100% of the time for six of eight measures of goodness-of-fit, and roughly 75% of the time for the remaining two measures. Indeed, given the lower number of parameters and the greater parsimony, we may well not wish to favour any distribution derived from a complex theory until it is shown to perform better than the lognormal. Moreover, the UNTB starts with assumptions that are known to be wrong: the competitive equivalence of species[6,8]. If the model is shown to have predictive power despite this, then we must question whether these competitive differences are important. But until there is a prediction of the UNTB that is strongly (that is, compared with a null hypothesis) and successfully tested, we have no evidence that these factors are unimportant, and much evidence that they are indeed significant. □

Received 18 October 2002; accepted 25 March 2003; doi:10.1038/nature01583.
Published online 13 April 2003.

1. Hubbell, S. P. *A Unified Theory of Biodiversity and Biogeography* (Princeton Univ. Press, Princeton, 2001).
2. Hubbell, S. P. Tree dispersion, abundance and diversity in a tropical dry forest. *Science* **203,** 1299–1309 (1979).
3. Hubbell, S. P. & Foster, R. B. in *Community Ecology* (eds Diamond, J. & Case, T.) 314–328 (Harper & Row, Cambridge, Massachusetts, 1975).
4. Bell, G. Neutral macroecology. *Science* **293,** 2413–2418 (2001).
5. de Mazancourt, C. Consequences of community drift. *Science* **293,** 1772 (2001).
6. Enquist, B. J., Sanderson, J. & Weiser, M. D. Modeling macroscopic patterns in ecology. *Science* **295,** 1835–1837 (2002).
7. Levine, J. M. Species diversity and relative abundance in metacommunities. *Trends Ecol. Evol.* **17,** 99–100 (2002).
8. Abrams, P. A. A world without competition. *Nature* **412,** 858–859 (2001).
9. Clark, A. Macroecology comes of age. *Trends Ecol. Evol.* **17,** 352–353 (2002).
10. Condit, R. *et al.* Beta-diversity in tropical forest trees. *Science* **295,** 666–669 (2002).
11. Terborgh, J., Foster, R. B. & Nunez, P. V. Tropical tree communities: a test of the nonequilibrium hypothesis. *Ecology* **77,** 561–567 (1996).
12. Yu, D. W., Terborgh, J. & Potts, M. D. Can high tree species richness be explained by Hubbell's null model? *Ecol. Lett.* **1,** 193–199 (1998).

13. Price, J., Droege, S. & Price, A. *The Summer Atlas of North American Birds* (Academic, San Diego, 1995).
14. Robbins, C. S., Bystrak, D. & Geissler, P. H. *The Breeding Bird Survey: Its First Fifteen Years, 1965–1979* (US Department of the Interior, Fish and Wildlife Service, Washington DC, 1986).
15. Sauer, J. R., Hines, J. E. & Fallon, J. *The North American Breeding Bird Survey, Results and Analysis 1966–2000* Version 2001.2 (USGS, Patuxent Wildlife Research Center, Laurel, Maryland, 2001); available at ⟨ftp://pwrcftpr.er.usgs.gov/mp/bbs/DataFiles/⟩.
16. Pyke, C. R., Condit, R., Aguilar, S. & Lao, S. Floristic composition across a climatic gradient in a neotropical lowland forest. *J. Veg. Sci.* **12,** 553–566 (2001).
17. Burslem, D., Whitmore, T. & Brown, G. Short-term effects of cyclone impact and long-term recovery of tropical rain forest on Kolombangara, Solomon Islands. *J. Ecol.* **88,** 1063–1078 (2000).
18. Laurance, W., Ferreira, L., Rankin-De Merona, J. & Laurance, S. Rain forest fragmentation and the dynamics of Amazonian tree communities. *Ecology* **79,** 2032–2040 (1998).
19. Gaston, K. J. & Blackburn, T. M. *Pattern and Process in Macroecology* (Blackwell Science, Oxford, UK, 2000).
20. Tokeshi, M. Species abundance patterns and community structure. *Adv. Ecol. Res.* **24,** 111–186 (1993).
21. Pielou, E. C. *Mathematical Ecology* (John Wiley & Sons, New York, 1977).
22. Harte, J., Kinzig, A. P. & Green, J. Self-similarity in the distribution and abundance of species. *Science* **284,** 334–336 (1999).
23. Dewdney, A. K. A dynamical model of communities and a new species-abundance distribution. *Biol. Bull.* **198,** 152–165 (2000).
24. Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, UK, 1983).
25. McGill, B. J. Strong and weak tests of macroecological theory. *Oikos* (in the press).
26. Grimmett, G. R. & Stirzaker, D. R. *Probability and Random Processes* (Clarendon, Oxford, UK, 1992).
27. May, R. M. in *Ecology and Evolution of Communities* (eds Cody, M. L. & Diamond, J. M.) 81–120 (Belknap, Harvard Univ. Press, Cambridge, Massachusetts, 1975).
28. Hilborn, R. & Mangel, M. *The Ecological Detective: Confronting Models with Data* (eds Levin, S. A. & Horn, H. S.) (Princeton Univ. Press, Princeton, New Jersey, 1997).

Supplementary Information accompanies the paper on *Nature*'s website (**http://www.nature.com/nature**).

Acknowledgements I thank the numerous volunteers who spent many hours collecting the BBS data, and the professionals at Patuxent Wildlife Research Center, who have put the data into an accessible, high-quality format, as well as those who collected and published the enormously valuable BCI data set. I also thank M. Rosenzweig, B. Enquist and the students in their laboratories for stimulating my thinking in this area. The National Science Foundation provided funding for this research. B. Walsh was a useful sounding board on statistics. Finally, I thank S. Marx, M. Weiser, D. Faulk and W. Turner for suggestions on an earlier draft of this paper, especially M. Weiser for his clarity on the importance of stopping rules, and S. Marx for greatly increasing the clarity of the paper.

Competing interests statement The author declares that he has no competing financial interests.

Correspondence and requests for materials should be addressed to the author (e-mail: mail@brianmcgill.org).

Ecological interference between fatal diseases

P. Rohani*, C. J. Green†, N. B. Mantilla-Beniers† & B. T. Grenfell†

* *Institute of Ecology, University of Georgia, Athens, Georgia 30602-2202, USA*
† *Zoology Department, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK*

An important issue in population biology is the dynamic interaction between pathogens. Interest has focused mainly on the indirect interaction of pathogen strains, mediated by cross immunity[1–4]. However, a mechanism has recently been proposed for 'ecological interference' between pathogens through the removal of individuals from the susceptible pool after an acute infection. To explore this possibility, we have analysed and modelled historical measles and whooping cough records. Here we show that ecological interference is particularly strong when fatal infections permanently remove susceptibles. Disease interference has substantial dynamical consequences, making multi-annual outbreaks of different infections characteristically out of phase. So, when disease prevalence is high and is associated with significant mortality, it might be impossible to understand epidemic patterns by studying pathogens in isolation. This new ecological null model has important consequences for understanding the multi-strain dynamics of pathogens such as dengue and echoviruses.

The possibility that epidemics of unrelated pathogens might interact has been raised in the classical epidemiological literature[5], but has not been explained. Recently, a new mechanism has been proposed for negative ecological interference between pathogens through the temporary removal of susceptibles, arising from infection by a competing pathogen and the ensuing quarantine period[6]. Interference should be particularly apparent in the violent recurrent epidemics of strongly immunizing childhood infections such as measles and whooping cough. However, recent parallel records of the two infections in England and Wales show equivocal evidence for interference, partly because of the relatively low pathogenicity of the infections[6]. Here we test for interference in older data, collected when measles and whooping cough were significant killers[5]. Because a fatal infection involves the permanent removal of susceptibles, we would expect the imprint of interference to be particularly strong.

We begin by exploring the predictions of simple models, based on extensions of the classic one-disease seasonally forced SEIR (susceptible–exposed–infectious–removed) model[7–9], with two important biological refinements (see Methods). First, the model includes a convalescent class[6], within which disease-induced deaths can occur; and second, we model the dynamics of two diseases simultaneously, categorizing hosts according to infection history relative to each disease. We are interested primarily in evaluating the dynamical impact of quarantine and disease-induced mortality on the community of pathogens. As with single-disease models, the dynamics of this system are determined largely by the recruitment rate of susceptibles (that is, the population birth rate[7,8,10]; Fig. 1). Very low/high per-capita birth rates result in annual epidemics, with biennial dynamics observed for intermediate levels[8]. When there is a disease-related mortality rate, ρ, the window of biennial behaviour is progressively delayed, with the period-doubling bifurcation taking place at higher birth rates (Fig. 1a). This is because high mortality due to one infection in effect lowers the recruitment rate of susceptibles for the 'competing' disease. In general, the bifurcation structure of the model is dictated by the infection with the higher transmission rate—in this case, measles[6]. Whooping cough epidemics, which in isolation would be rigidly annual for all parameter combinations, now follow the same pattern as measles (Fig. 1b–d).

When epidemics are annual because of a low/high birth rate[8], seasonal forcing causes strong positive correlation between infection outbreaks. However, given biennial epidemics, measles and whooping cough outbreaks are negatively correlated (out of phase)—much more so than if their dynamics were independent[6]. We use this negative correlation between disease dynamics as an indicator of potential interference in data.

We test for interference effects in case fatality reports for measles and whooping cough from Aberdeen (1883–1900) and from 15 European cities in the years before (1904–1914) and after (1922–1932) the First World War (Fig. 2). These data encompass large demographic heterogeneities, both spatial (between cities) and temporal (in the different eras). There were also systematic declines in the measles- and whooping-cough-induced death rates between periods before and after the war, which, along with the wide range of birth rates, provide an excellent opportunity to test the interference hypothesis in different regions of parameter space.

Children are typically affected by many more microparasitic infections than just measles and whooping cough (mumps, rubella and chickenpox within the childhood infections alone). However, we have focused on these two infections, because interference is likely to be most pronounced when diseases have very similar mean ages at infection, as dictated by their basic reproductive ratio, R_0 (ref. 7). Of potential childhood infections prevalent in the eras